On Metric Ramsey-Type Phenomena

Manor Mendel, California Institute of Technology

Joint work with:

Yair Bartal, The Hebrew University

Nati Linial, The Hebrew University

Assaf Naor, Microsoft Research

Metric Ramsey Problem

The Philosophy of Ramsey Theory:

Every system contains large highly structured subsystem.

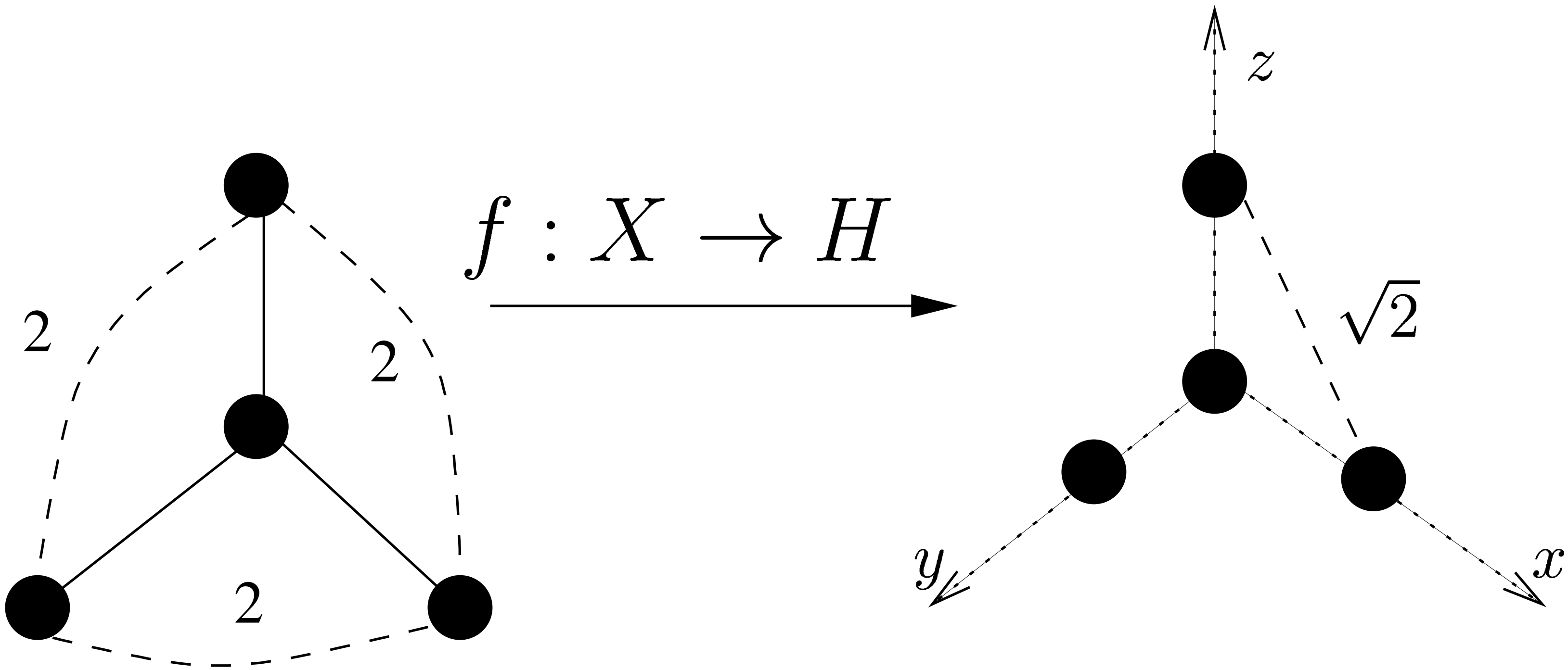
Metric Ramsey Problem:

Every finite metric space contains a *well-structured* subset of size $R(n)$.

Well-structured metric space:

Having low-distortion metric embedding in (say) Euclidean space.

Distortion



$X = K_{1,3}$ $\qquad$ $H = \ell_2^3$

dist$(f) = \sqrt{2}$, $\quad X$ $\sqrt{2}$-embeds in H.

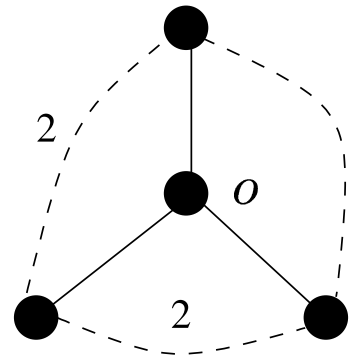
Metric Ramsey Formulation [BFM]

Given: $\alpha \geq 1$; X: n-point metric space.
What is the largest $Y \subset X$ such that
Y α-embeds in ℓ_2?

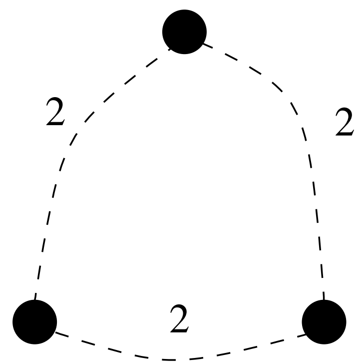
Def. $R_{\ell_2}(\alpha, n)$:
The largest m satisfying:
$\forall n$-point X, $\exists Y \subset X$ such that

1. Y α-embeds in ℓ_2.
2. $|Y| = m$.

(Prop. $R_{\ell_2}(1, n) \leq 3$,
so we will study $\alpha > 1$)



$K_{1,3}$ $(2/\sqrt{3})$-embeds in ℓ_2.



$(K_{1,3} \setminus \{o\})$ 1-embeds in ℓ_2

GA Motiv: Dvoretzky's Theorem [Dvo,Mil]

Thm.
$\forall \alpha > 1$, $\forall X$: n-dim' normed space,
$\exists Y \subset X$ subspace such that

1. Y α-embeds in ℓ_2.
2. dim$(Y) \geq C(\alpha) \log n$.

This bound is tight for the ℓ_∞ norm, for any $\alpha > 1$.

[Usually introduced via spheric sections of convex bodies]

- A cornerstone of modern convex geometry.
- Metric Ramsey problem is a metric analog of Dvoretzky's Theorem.

[BFM] result

Thm. [BFM]

1. $\forall \alpha > 1$, $R_{\ell_2}(\alpha, n) \geq C(\alpha) \log n$.
2. $\exists \alpha_0 \approx 1.023$, $R_{\ell_2}(\alpha_0, n) \leq 2 \log n + 3$.

These bounds are “similar to" Dvoretzky's Thm — Except: In the linear world, the $O(\log n)$ upper bound holds for any distortion $\alpha > 1$.

CS motiv: MTS, k-server

Online problems: k-server [MMS] , metrical task systems [BLS].

- A metric space X — part of their description.
- Cost = Sum of distances.
- An alg' for X induces alg's $\forall Y \subset X$.

Goal: Prove a lower bound $\forall$ n-point metric space.

Basic Technique: [KRR,BKRS,BBM]

1. Extract large $Y \subset X$, Y resembles a simple metric Z.
2. Prove a lower bound for Z.

Step 1. is a metric Ramsey problem.

Our Results

Thm.

1. $\forall \alpha \in (1,2)$,

$$c(\alpha) \log n \leq R_{\ell_2}(\alpha, n) \leq 2 \log n + C(\alpha).$$

2. $\forall \alpha > 2$, $\exists 0 < c(\alpha) \leq C(\alpha) < 1$,

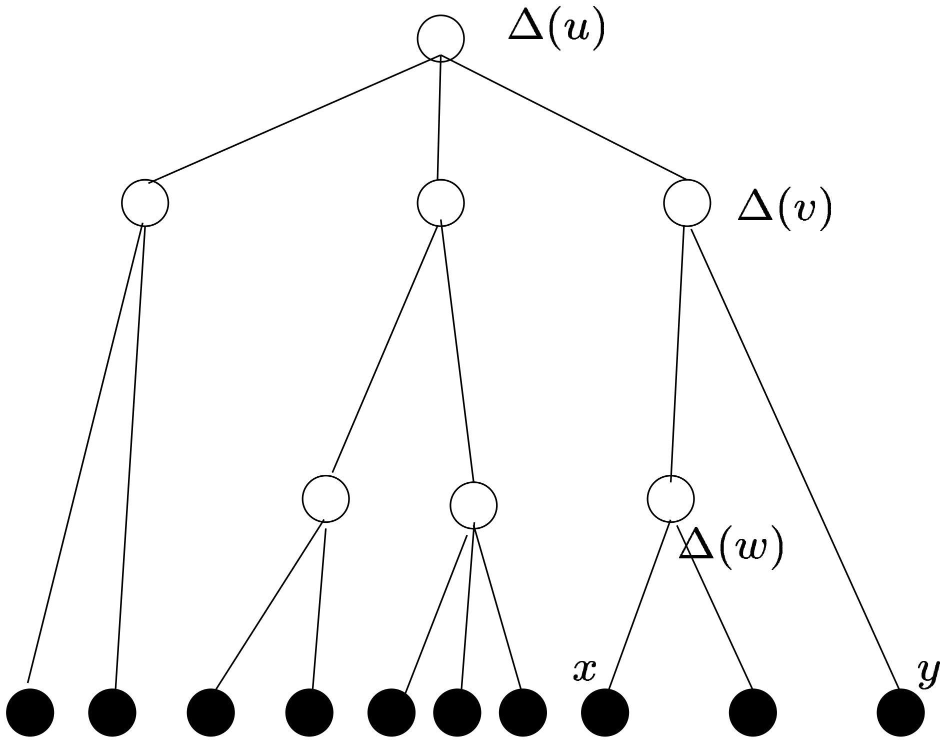
$$n^{c(\alpha)} \leq R_{\ell_2}(\alpha, n) \leq n^{C(\alpha)}.$$

3. $\forall \alpha > 2$ (large),

$$n^{1-\frac{c \log \alpha}{\alpha}} \leq R_{\ell_2}(\alpha, n) \leq n^{1-\frac{C}{\alpha}}.$$

Ultrametric / k-HST [B]

- Ultrametric = 1-HST.
- Observation: Ultrametrics 1-embed in ℓ_2.



$$\Delta(w) \leq \frac{\Delta(v)}{k} \leq \frac{\Delta(u)}{k^2}$$

$$d(x,y) = \Delta(v)$$

Example of k-HST

$R_{\mathbf{UM}}(\alpha, n) \geq n^{c(\alpha)}$, for $\alpha > 2$

An almost tight bound asymptotically in α:

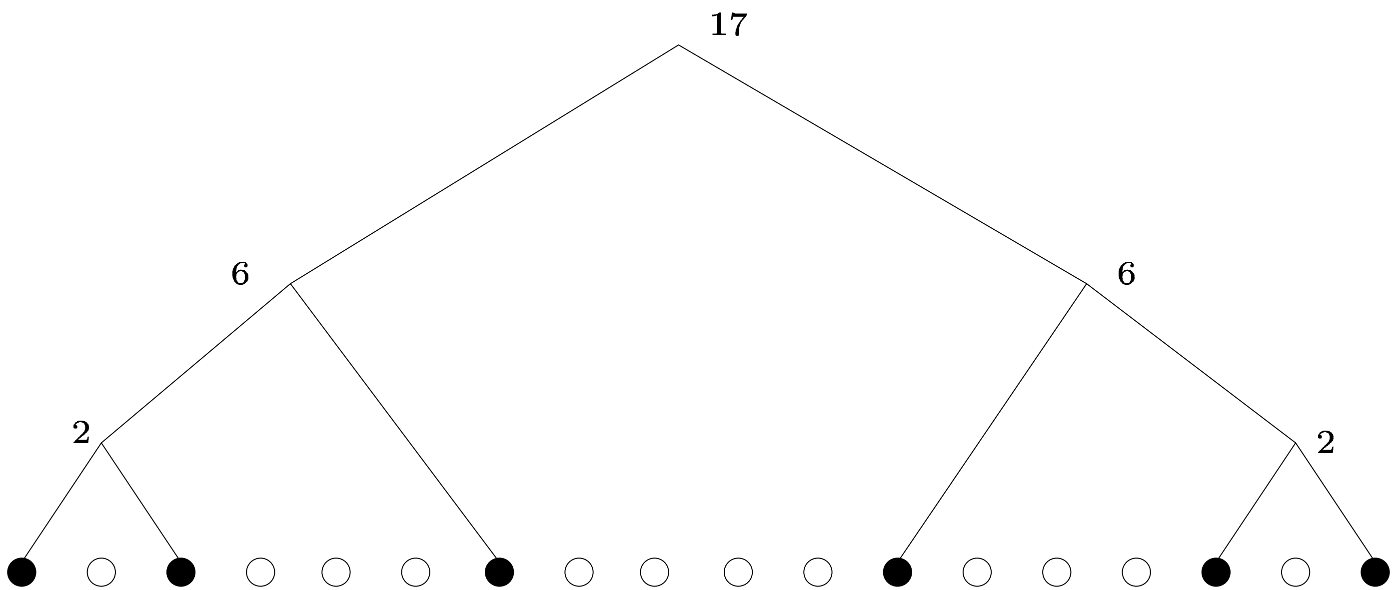
$$R_{\mathsf{UM}}(\alpha, n) \geq n^{1-\frac{c \log \alpha}{\alpha}}. \quad (1)$$

In other words,

$\forall \varepsilon > 0$, $\forall$ n-point X, $\exists Y \subset X$ s.t.

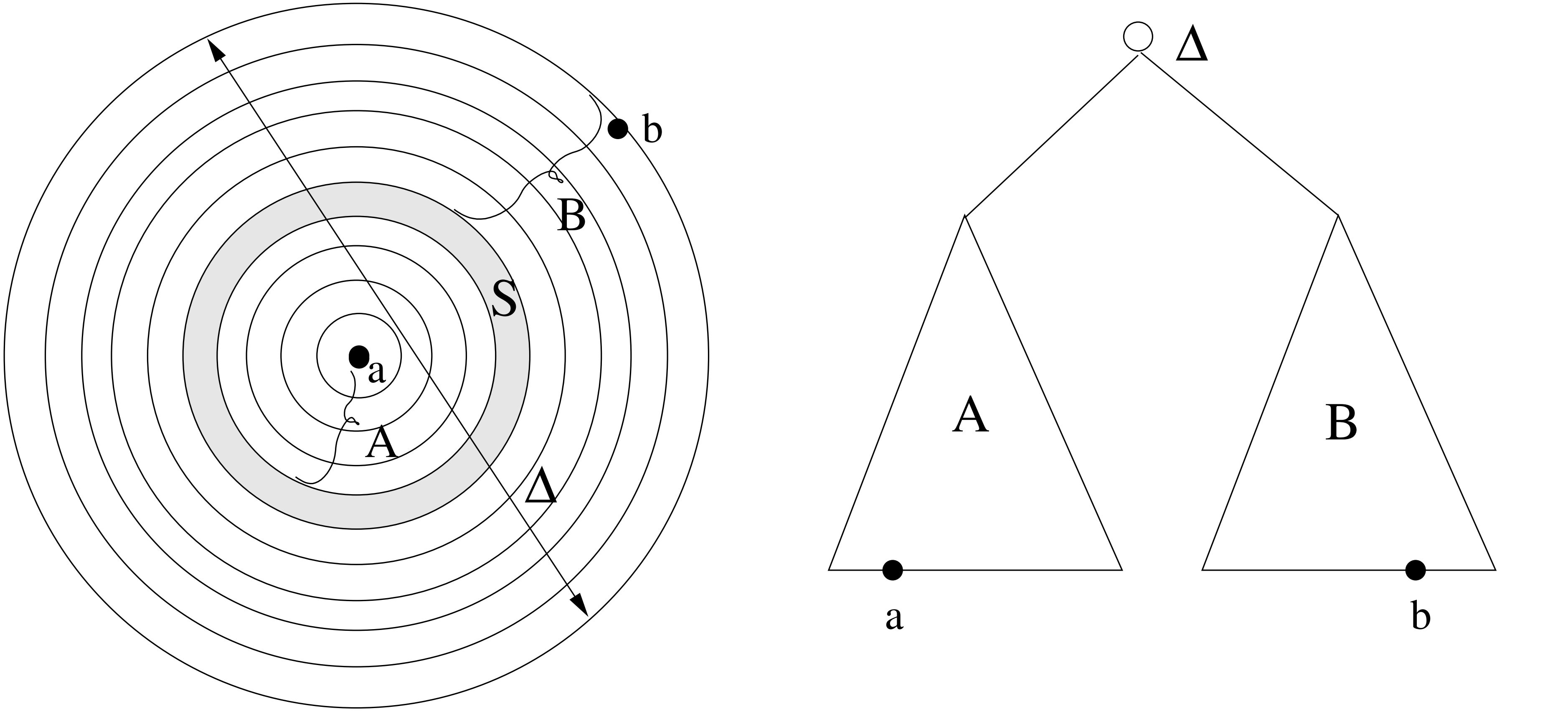
1. Y $O(\frac{\log(1/\varepsilon)}{\varepsilon})$-embeds in an ultrametric.
2. $|Y| \geq n^{1-\varepsilon}$.

A Toy Problem: Path Metric



- Extract a discrete version of the Cantor set.
- Obtain $2^{\log_3 n} = n^{\frac{\log 2}{\log 3}}$-point subset which 3 approximates UM (3-HST).
- Easy to see: This is essentially tight for UM.

Naïve Approach I [BBM]



Pick one shell S, and delete it.

Naïve Approach II

- Denote $|A| = \delta_i n$ points; $|B| = (1 - \delta_{i+1})n$.
- We want to prove $R_{\mathsf{UM}}(\alpha, n) \geq n^{\psi}$.
- By induction,

$$R_{\mathsf{UM}}(\alpha, n) \geq (\delta_i n)^{\psi} + ((1 - \delta_{i+1})n)^{\psi} = (\delta_i^{\psi} + (1 - \delta_{i+1})^{\psi})n^{\psi}$$

- It is sufficient to prove

$$\forall \frac{1}{n} \leq \delta_0 \leq \delta_1 \leq \cdots \delta_\alpha \leq 1 - \frac{1}{n} \quad \exists i, \quad \delta_i^{\psi} + (1 - \delta_{i+1})^{\psi} \geq 1.$$

- It is true for $\psi = 1/(\log n)^{2/\alpha}$.
- But it is not true for larger ψ.

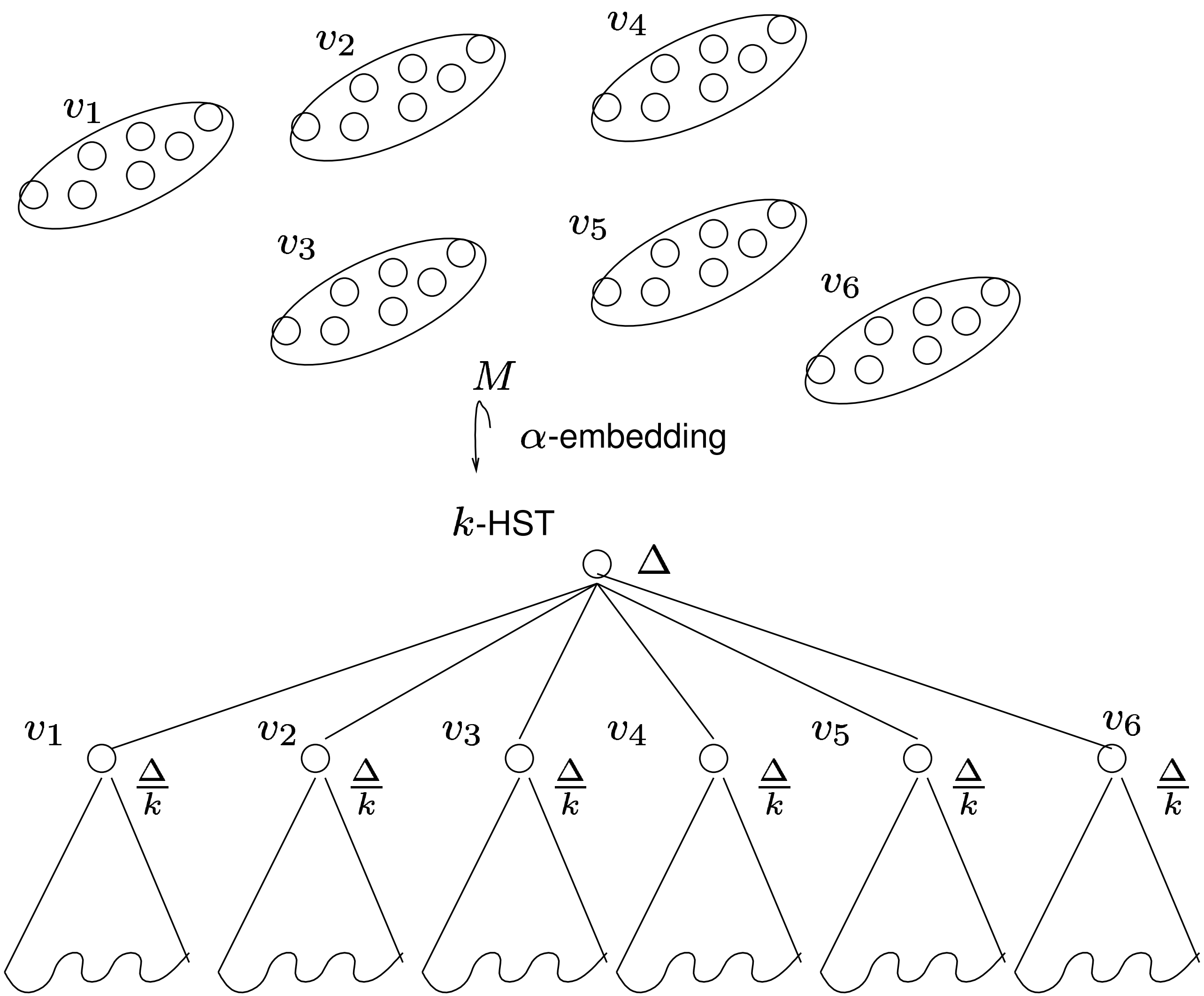
Our Construction

Key Idea:

Iteratively improve the embedding into an HST by deleting more points, using information from the previous embedding.

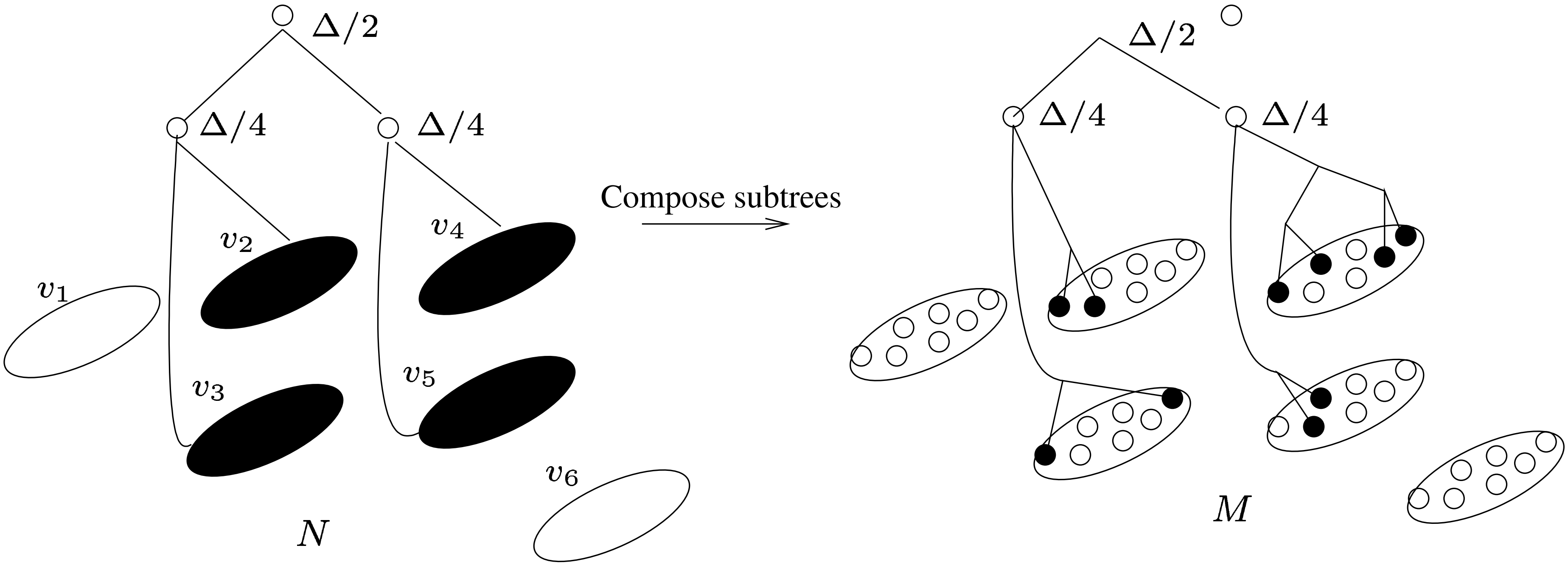
- Assume M is α-embedded in k-HST, and $k \geq 10\alpha$.
- Consider a vertex u and its children v_i in the HST.
- A natural metric N is defined on $\{v_i\}$.
- N approximates the distances in M (between subspaces).
- Solve a (weighted) Ramsey problem (into HSTs) for N, and concatenate the trees.

Composition I



We assume $k \geq 10\alpha$

Composition II



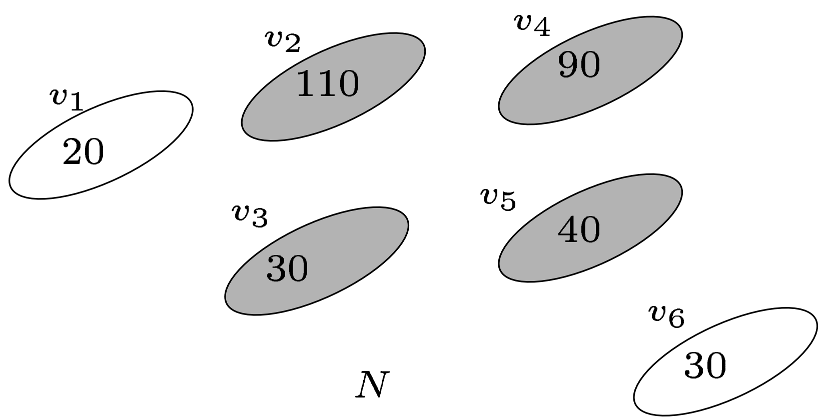
The Gain So Far

- A reduction from the Ramsey problem on M to a "weighted" Ramsey problem on N.
- Gain:
  - N has a controlled aspect ratio.
- Lose:
  - The weighted Ramsey problem is possibly more complicated.
- The trade-off is a progress!

Weighted Metric Ramsey Problem I

The high level Ramsey problem is “weighted": Find a subspace & largest ψ satisfying

$$110^\psi + 90^\psi + 30^\psi + 40^\psi \geq (110+90+30+40+20+30)^\psi.$$



Given a metric space N, distortion α. $\psi(\alpha, N)$ is the largest ψ s.t.:

$$\forall (w_u \geq 0)_{u \in N}\ \exists Y \subset N, s.t.$$

Y is α approximated by UM, and $\displaystyle\sum_{u \in Y} w_u^\psi \geq \Big(\sum_{u \in N} w_u\Big)^\psi.$

Weighted Metric Ramsey Problem II

Define

$$\psi(\alpha, \Phi) = \min\{\psi(\alpha, M);\ \text{a.r.}(M) \leq \Phi\}.$$

Lemma. $\psi(\alpha, \Phi) \geq 1 - c\frac{\log\log\Phi + \log\alpha}{\alpha}$.

- The proof is very technical with lots of details.
- The heart of the proof are [BBM]-like constructions, that consider both the aspect ratio and the weight.

One More Step I

In previous slides:

- Started with M_1, α_1-approx' by k_1-HST T_1, $k_1 \geq 10\alpha_1$.
- Found $M_2 \subset M_1$, α_2- approx' by 1-HST T_2 ($\alpha_2 \ll \alpha_1$).

Missing step:

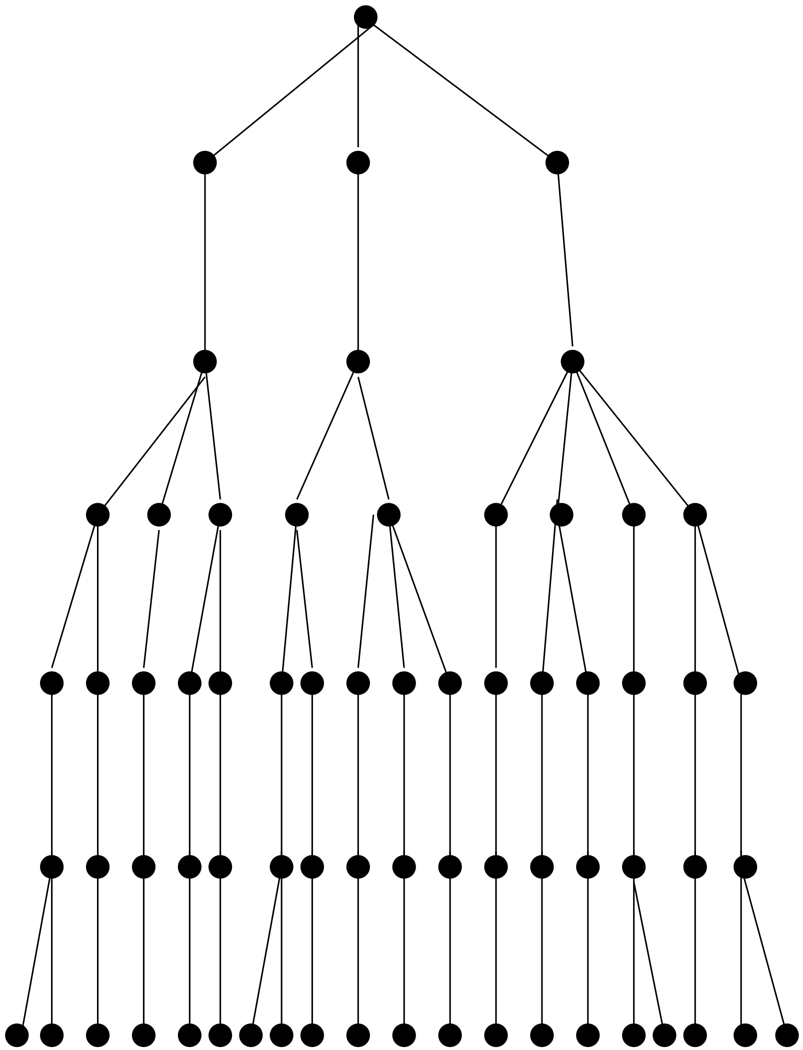
- Find $T_3 \subset T_2$, α_3 approx' by k_3-HST, $k_3 \geq 10\alpha_2\alpha_3$.
- Then $M_3 \subset M_2$ is $\alpha_2\alpha_3$ approximated by a k_3-HST.

One More Step II

The heart of the argument for the “missing step":

Def. A rooted tree is h-periodically sparse if on every vertical path every h-th vertex has only one child.

Lemma. Any n-leaf rooted tree contains h-periodically sparse subtree with $n^{\frac{h-1}{h}}$ leaves.



3 priodically sparse tree

The High Level Iteration

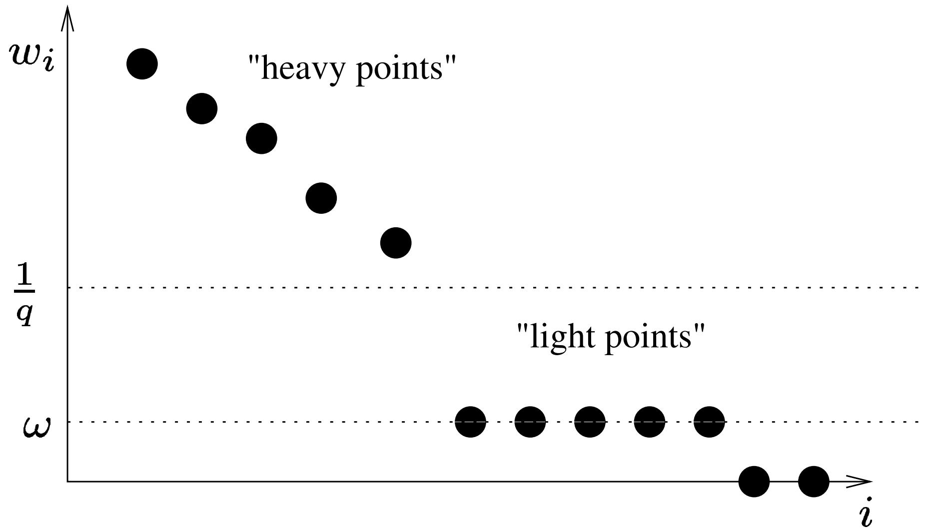
- Given Metric space M, distortion α.
- $n = |M|$, $\Phi =$ a.r.(M).
- M has trivial Φ-embedding into UM.
- Set $\beta_0 = \Phi$, $M_0 = M$.
- Iterate on i, until $\beta_i \leq \alpha$.
  - Assuming M_{i-1} has β_{i-1} embedding in UM.
  - Find $M_i \subset M_{i-1}$ having $\beta_i = \beta_{i-1}/2$ embedding in UM, and $|M_i| \geq |M_{i-1}|^{1-c\frac{\log \beta_{i-1}}{\beta_{i-1}}}$.
- At the end, M_{i_*} is $\beta_{i_*} \leq \alpha$ embedded in UM, and

$$|M_{i_*}| \geq n^{(1-c\frac{\log \alpha}{\alpha})(1-c\frac{\log 2\alpha}{2\alpha})(1-c\frac{\log 4\alpha}{4\alpha})\cdots} \geq n^{1-6c\frac{\log \alpha}{\alpha}}.$$

Proof of $\psi(\alpha, \Phi) \geq 1 - c\frac{\log\log\Phi + \log\alpha}{\alpha}$.

First step: Simplify the the weights.

Def. A sequence of weights $(w_i)_{i\in\mathbb{N}}$, $\sum_i w_i = 1$ is q-decomposable if exists $\omega > 0$ such that:



Lemma. By loosing a factor of $1 - \frac{c\log\log q}{\log q}$ in the exponent, we may assume the weights are q-decomposable.

Proof of $\psi(\alpha, \Phi) \geq 1 - c\frac{\log\log\Phi + \log\alpha}{\alpha}$.

Second Step: Apply a [BBM] like construction.
Case analysis at each level of the recursion:

- If there are two “heavy" points i, j with $d(i, j) \geq$ diam$/4$:
  - Apply the [BBM] argument with i, j as endpoints. In this case we can take $\psi = 1/(\log q)^{4/\alpha} \approx 1 - c\frac{\log\log q}{\alpha}$.
- Otherwise, there are two “light" points i, j, with $d(i, j) \geq$ diam$/4$, and i, j are far from the “heavy points":
  - Again, apply [BBM]-like procedure:
  - All the relevant shells contains only light points — so unweighted.
  - The inner ball has a.r. smaller by factor of 2 — enables induction on the aspect ratio.

Summary

- Comprehensive study of the metric Ramsey problem.
- Improved lower bounds for some online problems.
- Possible applications for clustering should be further studied.
- More results:
  - The metric Ramsey problem for special metrics:
    - Expanders, discrete cubes, high-girth graphs.
  - Embedding in low dimensional normed space.
  - Isometric embedding in ℓ_p and uniformly convex normed spaces.
  - Dichotomic metric Ramsey problems.